Predicted Binding Sites for the Regulatory Small Vault RNAs on Messenger RNAs of Selected Genes Relating to Cancer, Multi-drug Resistance, and Inflammation

Craig Jackson

Predicted Binding Sites for the Regulatory Small Vault RNAs on Messenger RNAs of Selected Genes Relating to Cancer, Multi-drug Resistance, and Inflammation

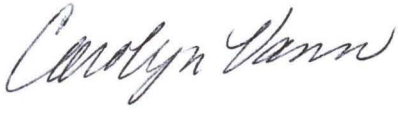
An Honors Thesis (HONRS 499)

By

Craig Jackson

Thesis Advisor

Dr. Carolyn Vann



Ball State University

Muncie, Indiana

May 2010

Expected Date of Graduation
May 2010

Abstract

Vaults are large ribonucleoprotein particles believed to be involved in multidrug resistance and intracellular transport. The vault complex consists of three proteins and non-coding vault RNAs. It has been shown that the non-coding vault RNA encodes regulatory small vault RNAs (svRNAs). These svRNAs associate with the RNA-induced silencing complex and regulate gene expression similarly to microRNAs. It is unknown which genes the svRNAs regulate, but they are thought to regulate genes relating to multidrug resistance and intracellular antigen transport. I have selected several genes of interest relating to cancer, multidrug resistance, inflammation, and the autoimmune response to predict whether the svRNAs regulate these genes. Since the svRNAs regulate genes similarly to microRNA, I used microRNA target prediction tools to find potential functional binding sites for the svRNAs in the 3' untranslated regions of the selected gene messenger RNAs. I found several target sites that have a very high potential to be functional binding sites for the svRNAs. These results can be used to conduct experiments to verify that the svRNAs bind to the predicted target sites and regulate the expression of the targeted gene.

Acknowledgements

Dr. Carolyn Vann – Project guidance, Helping me understand many biological concepts, Gene selection

Emmanuel Kwame Teye – Gene selection

Dr. Helena Persson – RNA alignment help and software recommendations

Table of Contents

Introduction 1

MicroRNAs 1

Vaults 3

Selected Gene Messenger RNAs 6

PTGS-1 (COX-1) 6

PTGS-2 (COX-2) 6

TEP1 6

NF- κB 7

NLRP1 7

PRKACB 8

IL10 (Interleukin 10) 8

CXCL14 8

ABCC2 9

ABCC4 10

ABCB1 (P-glycoprotein) 10

ABCG2 10

Methods 11

Results 14

PTGS-1 14

PTGS-1 Position 5005 Seed Match 14

PTGS-1 Position 4771 Seed Match 15

PTGS-1 Position 4325 Seed Match 15

PTGS-1 Position 4997 Seed Match 16

PTGS-1 Position 2751 Seed Match 16

PTGS-1 Position 2582 Seed Match 17

PTGS-1 Position 4898 Seed Match 17

PTGS-1 Position 3565 Seed Match 18

PTGS-1 Position 3667 Seed Match 18

PTGS-1 Result Remarks 19

PTGS-2 20

PTGS-2 Position 4052 Seed Match 20

PTGS-2 Position 3954 Seed Match 20

PTGS-2 Position 2191 Seed Match 21

PTGS-2 Result Remarks 21

TEP1 22

TEP1 Position 9230 Seed Match 22

TEP1 Position 9274 Seed Match 23

TEP1 Position 9140 Seed Match 23

TEP1 Position 10101 Seed Match 24
TEP1 Position 9373 Seed Match 24
TEP1 Position 8618 Seed Match 25
TEP1 Position 8700 Seed Match 25
TEP1 Position 9694 Seed Match 26
TEP1 Position 9662 Seed Match 26
TEP1 Position 10395 Seed Match 27
TEP1 Result Remarks 27
NFKB1 28
NFKB1 Position 3476 Seed Match 28
NFKB1 Position 3643 Seed Match 28
NFKB Result Remarks 29
NFKB2 29
RELA 30
RELA Position 2229 Seed Match 30
RELA Position 2104 Seed Match 30
RELA Position 2483 Seed Match 31
RELA Result Remarks 31
RELB 32
RELB Position 1970 Seed Match 32

RELB Position 2235 Seed Match 32

RELB Result Remarks 33

REL 33

REL Position 2561 Seed Match 33

REL Position 2536 Seed Match 34

REL Position 2395 Seed Match 34

REL Result Remarks 35

NLRP1 35

NLRP1 Position 5158 Seed Match 36

NLRP1 Result Remarks 36

PRKACB 37

PRKACB Position 3506 Seed Match 37

PRKACB Position 2382 Seed Match 38

PRKACB Position 1988 Seed Match 38

PRKACB Position 3848 Seed Match 39

PRKACB Position 3445 Seed Match 39

PRKACB Position 3641 Seed Match 40

PRKACB Position 1309 Seed Match 40

PRKACB Result Remarks 41

IL10 (Interleukin 10) 41

IL10 Position 1165 Seed Match 42
IL10 Position 981 Seed Match 42
IL10 Position 1287 Seed Match 43
IL10 Position 1023 Seed Match 43
IL10 Position 680 Seed Match 44
IL10 Position 764 Seed Match 44
IL10 Position 1368 Seed Match 45
IL10 Position 1453 Seed Match 45
IL10 Position 1428 Seed Match 46
IL10 Result Remarks 46
CXCL14 47
CXCL14 Position 1144 Seed Match 47
CXCL14 Position 1174 Seed Match 47
CXCL14 Position 1276 Seed Match 48
CXCL14 Result Remarks 48
ABCC4 49
ABCC4 Position 5302 Seed Match 49
ABCC4 Position 5190 Seed Match 49
ABCC4 Position 4255 Seed Match 50
ABCC4 Position 5005 Seed Match 50

ABCC4 Result Remarks 51

ABCG2 51

ABCG2 Position 3567 Seed Match 52

ABCG2 Position 3555 Seed Match 52

ABCG2 Position 4020 Seed Match 53

ABCG2 Position 3970 Seed Match 53

ABCG2 Result Remarks 54

Introduction

MicroRNAs

MicroRNAs (miRNAs) are highly conserved small non-coding RNAs, usually about 22 nucleotides in length, which regulate gene expression. This regulation takes place mostly by targeting the 3' untranslated region (3' UTR) of messenger RNAs (mRNAs). miRNAs guide the RNA-induced silencing complex (RISC) to target sites in the mRNA that have reverse complementary sequences with the miRNA (5' end of miRNA is aligned with 3' end of mRNA, and vice versa) (Hofacker, 2007). When the RISC, bound with a miRNA, finds a complementary match, it suppresses the translation of the target mRNA, or degrades the mRNA with the use of ribonuclease. Both of these actions prevent the creation of the protein that the target mRNA codes for.

The complementarity between miRNAs and target sites on mRNAs are far from perfect matches though. The defining part of a miRNA-mRNA match is a short seed region of 6-8 nucleotide bases. This seed region is located on the 5' end of the miRNA at nucleotide positions 2 through 9. The seed region of the miRNA is sometimes the only part of the sequence that is perfectly base-paired with the mRNA. As such, these seed matches are used by miRNA target prediction algorithms to find potential binding sites on mRNAs (Hofacker, 2007). Because using such a short seed region to find target sites will produce many false positive hits, the miRNA target prediction algorithms must use some other data to produce reliable predictions. Many methods, such as PicTar and TargetScan, rely heavily on evolutionary conservation of the target site to determine the likelihood of a positive match. But evolutionary conservation is not useful for species-specific miRNAs, and it does not address the mechanism by which the miRNAs find their target mRNAs in the cell (Hofacker, 2007).

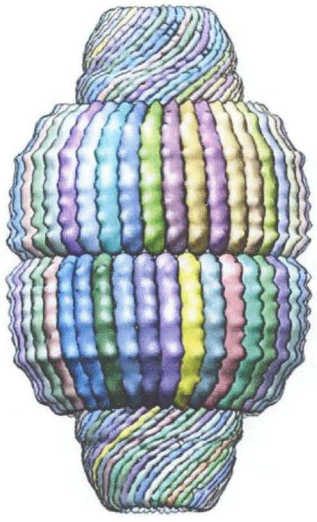
Kertesz *et al.* developed a new algorithm, called Probability of Interaction by Target Accessibility (PITA), to predict miRNA targets using free energy calculations of the miRNA-mRNA

bindings instead of evolutionary conservation. These same free energy calculations are used by several algorithms to predict the minimum free energy (MFE) RNA secondary structures. According to the principle of minimum energy, which is a restatement of the second law of thermodynamics, RNA will form secondary structures that minimize the free energy of the system. The PITA algorithm calculates three different free energies at a target mRNA site to predict the site accessibility. Previous methods for miRNA target prediction scored interactions based on only the free energy gained by binding of the miRNA to the mRNA target sequence. This free energy measurement is called ΔG_{duplex}. The PITA first calculates this ΔG_{duplex}, but does not stop there like previous algorithms. Next, the PITA algorithm calculated the free energy lost by unpairing the mRNA target site from its RNA secondary structure, called ΔG_{open}. The PITA algorithm uses the free-energy secondary structure prediction algorithm of (Zuker and Stiegler, 1981) from the Vienna RNA Package to predict the mRNA's secondary structure, which it then uses to find ΔG_{open}. The PITA algorithm then takes the difference between ΔG_{duplex} and ΔG_{open} to assign an energy-based score, called $\Delta\Delta G$, for the miRNA-mRNA association. Kertesz *et al.* found that their PITA algorithm predicted miRNA targets better than all of the previous miRNA target prediction algorithms (most of which use evolutionary conservation filters). The integration of site accessibility in miRNA target prediction represents a measurable improvement and a simplification over existing methods (Kertesz *et al.* 2007).

Genes were once thought of almost exclusively as instructions for building proteins, but now scientists' perception of them is changing. Genes are now seen as RNA factories, not just protein instruction sets ("Really New Advances," 2007). In fact, genes for protein may actually be in the minority. According to Isidore Regoutsos, IBM's genome-miner in chief, there may as many as 37,000 miRNAs in humans, while there are 21,000 protein-encoding genes in humans. It is now thought that miRNAs regulate expression of the majority of genes in the human genome (Peter,

2010). The existence of miRNAs may also explain how some organisms are more complex than others, even though most animals have around 20,000 genes. If complexity was just related to the number of genes in an organism, then we would not be able to explain the increased complexity in some organisms, such as humans. But with the new knowledge of RNA gene regulation, we may be able to begin to explain the differences in organism complexity based on the role RNA plays in gene regulation in the more complex organisms.

Vaults



Vaults are large ribonucleoprotein particles implicated in multidrug resistance and intracellular transport. The vault complex is the largest ribonucleoprotein particle described to date, with a molecular mass of 13 MDa, approximately three times the size of a ribosome (van Zon *et al.*, 2001). Vaults are conserved organelles, found in many diverse eukaryotic cells, including in mammals, birds, fish, echinoids, and slime molds (Persson *et al.* 2009). Discovered in 1986, the structures were named vaults because their morphology resembled the arches of vaulted ceilings in cathedrals. Vaults form hollow barrel-like structures that are symmetrical and have an invaginated waist and two protruding caps at each end (van Zon *et al*, 2001). Three proteins make up the vault

complex – the major vault protein (MVP), poly(ADP-ribose) polymerase 4 (PARP4/vPARP), and telomerase-associated protein 1 (TEP1). The barrel-like waist of the vault particle is made up of 96 MVP molecules. The protruding caps on the end of vault particle are made up of 8 PARP4 molecules and 2 TEP1 molecules (van Zon *et al*, 2001). The vault complex also contains non-coding vault RNA. While the three vault proteins are structural components of the vault complex, the vault RNA instead plays a functional role.

Persson *et al.* found that vault RNAs could theoretically fold into structures resembling microRNA precursors, so they decided to investigate whether the vault RNA could be a new source of small regulatory RNAs. They found that multiple regulatory small vault RNAs (svRNAs) were encoded by the vault RNA. The svRNAs were produced by a Dicer-dependent but Drosha-independent mechanism (Persson *et al.*, 2009). Dicer is a component of the microRNA pathway, so Persson *et al.* tested to see whether the svRNAs participated in the RISC complex like microRNAs. They found that the svRNAs could indeed participate in the RISC complex, regulating gene expression similar to miRNAs. Approximately 80% of vault RNA is free in the cytosol, suggesting that the svRNAs could be processed from the accessible cytosolic vault RNA, without the presence of the vault organelle.

Vault particles have been found to be up-regulated in multidrug-resistant cancer cell lines (Kickhoefer, 1998). Multidrug resistance is the major cause of chemotherapy failure in cancer treatment, and is an important current field of research. Multidrug-resistant cancer cells frequently overexpress a protein that was originally named the Lung Resistance-related Protein. Investigation of this protein revealed it was the human homologue of the Major Vault Protein (Scheffer *et al.*, 1995). The MVP has since been shown to be overexpressed in many multidrug-resistant tumor cell lines, including SW1573/2R120 non-small cell lung cancer, GLC4/ADR small cell lung cancer, MCF-7/MR breast cancer, and 8226/MR20 myeloma (Kickhoefer, 1998). It has been shown that

the overexpression of the MVP alone is not enough to confer multi-drug resistance, which is not unexpected since MVP comprises only 70% of the vault particle. The minor vault proteins and the vault RNA (including the regulatory svRNAs) could also be required for drug resistance to occur.

Selected Gene Messenger RNAs

PTGS-1 (COX-1) - Homo sapiens prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase) (PTGS1), transcript variant 1, mRNA (NCBI Accession: NM_000962.2)

Prostaglandin-endoperoxide synthase, also known as Cyclooxygenase, is the central enzyme in the biosynthesis of prostaglandins. PTGS-1 encodes the first isozyme of Cyclooxygenase. Prostaglandins play an important role in inflammation, as well as other functions in animals. Because of its large role in inflammation, I chose to predict svRNA target sites in the 3' UTR of the PTGS-1 gene.

PTGS-2 (COX-2) - Homo sapiens prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (PTGS2), mRNA (NCBI Accession: NM_000963.2)

PTGS-2 encodes the second isozyme of Cyclooxygenase. It differs from the first isozyme because it is an inducible enzyme and is regulated by specific stimulatory events. PTGS-2 is upregulated in many cancers, and can stimulate cancer progression. Because of its role in inflammation and its link to cancer, I chose to predict svRNA target sites in the 3' UTR of the PTGS-2 gene.

TEP1 - Homo sapiens telomerase-associated protein 1 (TEP1), mRNA (NCBI Accession: NM_007110.4)

TEP1 is a minor vault protein that forms part of the vault structure, at the end caps of the vault. TEP1 is required for a stable association of the vault RNA with the vault particle, and helps recruit free vault RNAs to the vault particle (Kickhoefer *et al*, 2001). Therefore, target sites for the

small vault RNAs may exist in the TEP1 gene, in order to recruit the vault RNA into the vault particle. For this reason, I chose to predict svRNA target sites in the 3' UTR of the TEP1 gene.

NF- κB - Nuclear factor kappa-light-chain-enhancer of activated B cells

NF-κB is a protein complex that controls the transcription of DNA. NF-κB regulates many genes involved in inflammation and the immune system. Incorrect regulation of NF-κB has been linked to cancer, inflammatory and autoimmune diseases, and improper immune development. NF-κB is also a regulator of genes that control cell proliferation and cell death, protecting cells from conditions that cause apoptosis. In tumor cells, NF-κB is active due to mutations in genes that encode NF-κB or genes that control NF-κB activity. This causes the tumor cells to maintain cell proliferation, preventing apoptosis and allowing the cancer to progress. There are five proteins in the mammalian NF-κB complex: NF-κB1, NF-κB2, RelA, RelB, and c-Rel. Because of the NF-κB complex's role in inflammation, the immune system, and the proliferation of cancer cells, I chose to predict svRNA target sites in the 3' UTR regions of the 5 genes that encode the 5 NF-κB proteins. These genes are NFKB1 (NCBI Accession: NM_003998.3), NFKB2 (NCBI Accession: NM_001077494.1), RELA (NCBI Accession: NM_021975.3), RELB (NCBI Accession: NM_006509.2), and REL (NCBI Accession: NM_002908.2).

NLRP1 - Homo sapiens NLR family, pyrin domain containing 1 (NLRP1), transcript variant 1, mRNA (NCBI Accession: NM_033004.3)

The NLRP1 gene encodes a member of the Ced-4 family of apoptosis proteins. Ced-family members are known to be a key mediator of programmed cell death. Overexpression of the NLRP1 gene was demonstrated to induce apoptosis in cells. Because of its strong connection to apoptosis in cells, I chose to predict svRNA target sites in the 3' UTR region of NLRP1.

PRKACB - Homo sapiens protein kinase, cAMP-dependent, catalytic, beta (PRKACB), transcript variant 1, mRNA (Accession: NM_182948.2)

The PRKACB gene encodes the cAMP-dependent protein kinase catalytic subunit beta, which is an enzyme that is part of the cAMP-dependent pathway. Cyclic adenosine monophosphate, or cAMP, is a second messenger that is important in many cell functions, and is involved in the activation of protein kinases. Some research has suggested that deregulation of the cAMP pathways is linked to the growth of some cancers. PRKACB was found to be strongly upregulated in MCF-7 cancer cells, and was within the top 100 predicted target sites of svRNAb (Persson *et al*, 2009). Because of its link to cancer growth and its high svRNAb target site prediction, I chose to predict svRNA target sites in the 3' UTR region of the PRKACB gene.

IL10 (Interleukin 10) - Homo sapiens interleukin 10 (IL10), mRNA (Accession: NM_000572.2)

The IL10 gene encodes the protein Interleukin 10, which is an anti-inflammatory cytokine. Interleukin 10 has effects in immunoregulation and inflammation. It also enhances cell survival, proliferation, and antibody production. Interleukin 10 can inhibit the synthesis of many pro-inflammatory cytokines, and can block activity in the NF-κB complex. IL10 was found to be strongly upregulated in MCF-7 cancer cells (Persson *et al*, 2009). Because of its effects on immunoregulation and inflammation, as well as its potential role in cell proliferation and antibody production, I chose to predict svRNA target sites in the 3' UTR region of the IL10 gene.

CXCL14 - Homo sapiens chemokine (C-X-C motif) ligand 14 (CXCL14), mRNA (Accession: NM_004887.4)

The CXCL14 gene encodes the cytokine Chemokine ligand 14, as known as breast and kidney-expressed chemokine (BRAK). CXCL14 plays a role in inflammation and the immune system, by activating monocytes when in the presence of the inflammatory mediator prostaglandin-E2. CXCL14 was found to be strongly upregulated in MCF-7 cancer cells (Persson *et* al, 2009) Because of its effects on inflammation and the immune system, I chose to predict svRNA target sites in the 3' UTR region of the CXCL14 gene.

ATP-binding cassette transporters (ABC-transporters)

ABC-transporters are a superfamily of transmembrane proteins that utilize the energy of adenosine triphosphate (ATP) hydrolysis to translocate various substrates across cell membranes. They transport a wide variety of substrates across extracellular and intracellular membranes. Some of the ABC-transporters transport drugs across these membranes, and play a role in multidrug resistance. The remaining genes encode proteins of this superfamily that are potentially involved in drug transport and multidrug resistance. I chose to predict svRNA target sites on the genes because of their role in multidrug-resistance.

ABCC2 - Homo sapiens ATP-binding cassette, sub-family C (CFTR/MRP), member 2 (ABCC2), mRNA (Accession: NM_000392.3)

The ABCC2 gene encodes the Canalicular multispecific organic anion transporter 1 protein, which is an ABC-transporter protein. The ABCC2 protein is a member of the multidrug resistance protein (MRP) subfamily, which is involved in multidrug-resistance. The substrates that ABCC2 transport include anticancer drugs such as vinblastine. ABCC2 appears to be involved in drug resistance in mammalian cells.

ABCC4 - Homo sapiens ATP-binding cassette, sub-family C (CFTR/MRP), member 4 (ABCC4), transcript variant 1, mRNA (Accession: NM_005845.3)

The protein encoded by the ABCC4 gene is an ABC-transporter protein, and is a member of the multidrug resistance protein (MRP) subfamily, which is involved in multidrug resistance. ABCC4 transports organic anions, and has been found to transport cAMP (discussed above)(van Aubel *et al*, 2002).

ABCB1 (P-glycoprotein) - Homo sapiens ATP-binding cassette, sub-family B (MDR/TAP), member 1 (ABCB1), mRNA (Accession: NM_000927.3)

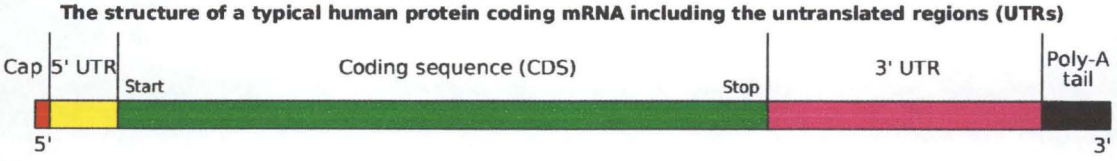
The ABCB1 gene encodes the ABC-transporter protein P-glycoprotein (PGP). PGP is a member of the multidrug resistance protein (MRP) subfamily, which is involved in multidrug resistance. PGP transport many substrates, including chemotherapeutic agents such as doxorubicin and vinblastine.

ABCG2 - Homo sapiens ATP-binding cassette, sub-family G (WHITE), member 2 (ABCG2), mRNA (Accession: NM_004827.2)

The ABCG2 gene encodes the ATP-binding cassette sub-family G member 2 protein, also known as the breast cancer resistance protein. ABCG2 functions as a xenobiotic transporter which may play a role in multidrug resistance.

Methods

After selecting the genes to do svRNA target predictions on, I used the National Center for Biotechnology Information (NCBI) Entrez Global Query Cross-Database Search System to find the gene sequences. Once I located the gene in the Entrez Gene database, I navigated to mRNA entry for the gene in the Entrez Nucleotide database. The Entrez Nucleotide database lists the entire mRNA transcript for the gene, but I only wanted to run predictions on the 3' UTR region of the gene. The Entrez Nucleotide entry for the gene lists the nucleotide positions of the coding sequence (CDS) of the gene. The coding sequence is the part of the gene's mRNA that is translated into the protein. The 3' UTR is located after the stop codon in the coding sequence on the 3' end of the mRNA. Knowing the ending nucleotide position of the coding sequence allowed me to find where the 3' UTR region of the mRNA started. I then selected the nucleotides from the end of the coding sequence until the end of the mRNA sequence from the FASTA formatted sequence on the NCBI website. This selected sequence includes the 3' UTR and the poly-A tail of the gene's mRNA.



Next, I used to the Segal Lab of Computational Biology's online microRNA prediction tool to predict svRNAa and svRNAb target sites in the selected genes' 3' UTR ("Online microRNA prediction tool"). The online microRNA prediction tool uses the Probability of Interaction by Target Accessibility algorithm developed by Kertesz *et al.* to find and score the target sites. I set the seed parameters for the prediction tool to find seed matches of 6 to 8 bases in length, with a single G:U wobble pair allowed, and no mismatches allowed. The recommended seed parameter settings are to consider seeds length 6 to 8 bases, with no mismatches, and up to one G:U wobble pair in 7-

or 8-mers. Since I allowed a wobble base pair in 6-mer seed matches in the parameters I used, there is a potential for false positives, and 6-mer seed matches with a wobble pair (6:0:1) in the results should be scrutinized more than the other matches. The more negative that the score assigned by the PITA algorithm (ΔΔG) is, the more likely that the microRNA (or svRNA in this case) will actually bind to the target site in natural circumstances. As a rough rule of thumb, sites with ΔΔG values below -10 are likely to be functional in endogenous microRNA (svRNA in this case) expression levels, although with very high expression levels, sites with ΔΔG above -10 may be functional as well. PITA's notation of the seed column results (X:Y:Z) represent the size of the seed (X), the number of mismatches (Y), and the number of G:U wobble pairs (Z).

Once I got the PITA seed match results for all of the selected genes from the online microRNA prediction tool, I began the svRNA-mRNA alignments for the target sites predicted by the PITA algorithm. First, I manually aligned the 22nt (svRNAa matches) or 23nt (svRNAb matches) sequences from the mRNA at the target site with the svRNAs in an one-to-one straight alignment (each nucleotide position in the svRNA sequence matched up with the same position in the mRNA sequence, with no gaps). Although this straight alignment is not usually how the svRNA would actually bind to the target RNA, it gives a general first look at the svRNA next to the target site on the mRNA.

Next I used the RNAfold web server (which uses the Zuker and Stiegler algorithm) from the Vienna RNA Websuite (Gruber *et al.*, 2008) to find the MFE binding configuration of the svRNA to the mRNA target site. The RNAfold program takes a single RNA sequence as its input and calculates the MFE secondary structure of that RNA sequence. Since I needed to find the MFE binding between two RNA sequences (the svRNA sequence and the mRNA target site sequence), I combined the two sequences together with four N's and gave folding constraints that required the seed regions of the two sequences to bind together. This forced the RNAfold algorithm to fold the

combined RNA sequence into a stem-loop structure around the four N's I inserted. This results in the MFE binding between the svRNA and the mRNA target site, which is most likely how they would bind in nature according to the minimum energy principle. A visual representation of the MFE predicted structure is output by RNAfold, and the base pairs are colored according to their base pair probability: 

.

Results

PTGS-1

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
PTSG-1 3' UTR	svRNAa	3070	6:0:0	-23.8	-7.72	-16.07
PTSG-1 3' UTR	svRNAb	2836	6:0:1	-17.1	-6.49	-10.6
PTSG-1 3' UTR	svRNAb	2390	8:0:1	-23.83	-14.21	-9.61
PTSG-1 3' UTR	svRNAb	3062	7:0:1	-14.1	-7.49	-6.6
PTSG-1 3' UTR	svRNAa	816	7:0:1	-19.2	-13.53	-5.66
PTSG-1 3' UTR	svRNAb	647	6:0:1	-16.4	-13.14	-3.25
PTSG-1 3' UTR	svRNAb	2963	6:0:0	-14.55	-12.47	-2.07
PTSG-1 3' UTR	svRNAa	1630	6:0:1	-12.4	-12.15	-0.24
PTSG-1 3' UTR	svRNAb	1732	6:0:1	-14.25	-16.44	2.19

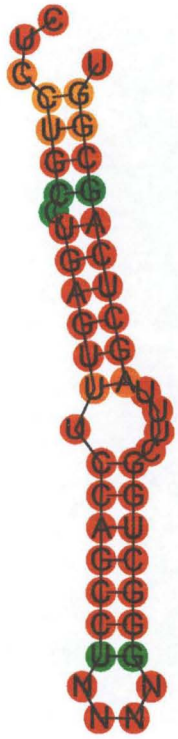
PTGS-1 Position 5005 Seed Match (3' UTR Position 3070)

One-to-One Alignment

```
                    5005
                     |
5' CUCCUGCCUGAGUUUCCAGCCU  3'  <- PTGS-1
    :||:       :    ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      5005
                       |
5' CUCCUGCCUGAGUU---UCCAGCCU  3'  <- PTGS-1
      |:|| ||||||    ||||||:
3'   UGGCG-ACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



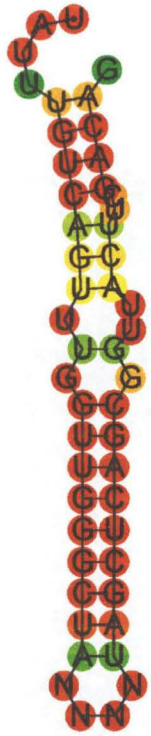
PTGS-1 Position 4771 Seed Match (3' UTR Position 2836)

One-to-One Alignment

```
                      4771
                      |
5' UAUUUGUCAGUUUGGUUGGGCUA  3'  <- PTGS-1
   :  |:|       : |:||:||||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                        4771
                        |
5' UAUUUGUC--AGUU-UGGUUGGGCUA  3'  <- PTGS-1
       ||||  |||  : |:||:||||
3'    GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



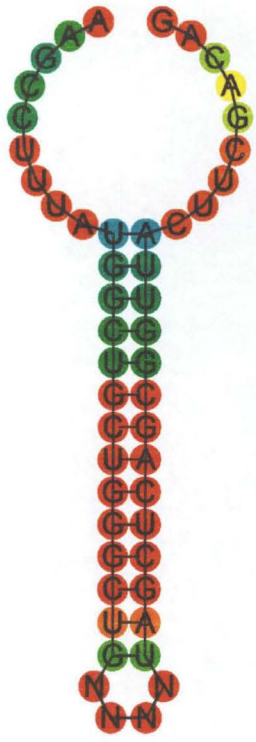
PTGS-1 Position 4325 Seed Match (3' UTR Position 2390)

One-to-One Alignment

```
                    4325
                    |
5' AAGCCUUUAUGGCUGCUGGGCUG  3'  <- PTGS-1
     | |    |::|:||||:|||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                    4325
                    |
5' AAGCCUUUAUGGCUGCUGGGCUG  3'  <- PTGS-1
            |::|:||||:|||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



PTGS-1 Position 4997 Seed Match (3' UTR Position 3062)

One-to-One Alignment

```
                    4997
                     |
5' GCCAUCGACUCCUGCCUGAGUUU  3'  <- PTGS-1
       : :| |  :  |||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      4997
                       |
5'  GCCAUCGACUC--CUGCCUGAGUUU  3'  <- PTGS-1
        ||||     |:| |||||:|
3'   GACAGCUUCAUUGGC-GACUCGAU  5'  <- svRNAb
```

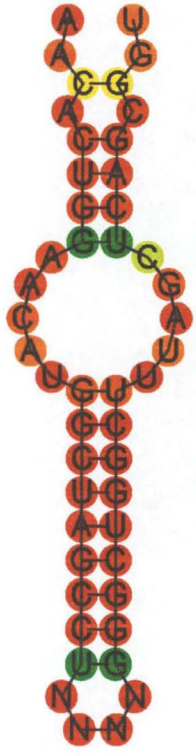
PTGS-1 Position 2751 Seed Match (3' UTR Position 816)

One-to-One Alignment

```
                  2751
                   |
5' AACACUGGAACAUGGCUAGCCU  3'  <- PTGS-1
   | | |||:   | :||:|||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                  2751
                   |
5' AACACUGGAACAUGGCUAGCCU  3'  <- PTGS-1
     | |||:     :||:|||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



PTGS-1 Position 2582 Seed Match (3' UTR Position 647)

One-to-One Alignment

```
                      2582
                       |
5' AUGUAGAGAGAACAGGUGGGCUG  3'  <- PTGS-1
    ||| ||:  ||| ||||:|||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      2582
                       |
5' AUGUAGAGAGAACAGGUGGGCUG  3'  <- PTGS-1
    ||| ||:  ||| ||||:|||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

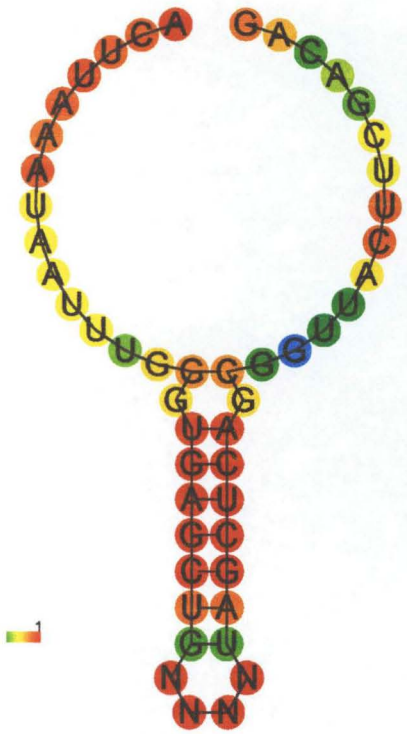
PTGS-1 Position 4898 Seed Match (3' UTR Position 2963)

One-to-One Alignment

```
                      4898
                       |
5' ACUUAAAUAAUUUGGGUGAGCUG  3'  <- PTGS-1
     |  |     : | ||||||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      4898
                       |
5' ACUUAAAUAAUUUGGGUGAGCUG  3'  <- PTGS-1
                  | ||||||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



PTGS-1 Position 3565 Seed Match (3' UTR Position 1630)

One-to-One Alignment

```
                    3565
                     |
5' AGACAGCCCUCCACUCCAGCUC  3'  <- PTGS-1
   |         :   |  |||||:|
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                          3565
                           |
5' AGACAGCC-CUCCACU----CCAGCUC  3'  <- PTGS-1
        :|| ||    ||    |||||:
3'      UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

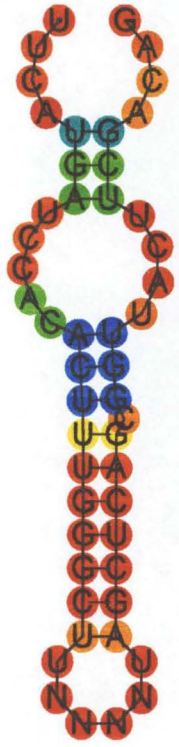
PTGS-1 Position 3667 Seed Match (3' UTR Position 1732)

One-to-One Alignment

```
                    3667
                     |
5' UUCAUGAUCCACACUUUGGGCUU  3'  <- PTGS-1
   :|  :||   |  | :||:|||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                     3667
                      |
5' UUCAUGAUCCACACU-UUGGGCUU  3'  <- PTGS-1
       :||      ||: :||:|||
3' GACAGCUUCAU-UGGCGACUCGAU  5'  <- svRNAb
```



PTGS-1 Result Remarks

PTGS-1 has some very promising seed matches that seem highly likely to be functional. The best scoring seed match, at position 5005, seems very promising with a very high ΔG_{duplex} value of -23.80 and a $\Delta\Delta G$ value of -16.07. In the predicted MFE binding structure, all but 5 of the svRNAa bases pair with the target site, and all the base pairs have very high base pair probability.

The second best scoring seed match, at position 4771, is also a good potential to be a functional match with $\Delta\Delta G$ value of -10.6 and ΔG_{duplex} value of -17.10. It is a 6-mer seed match with a wobble base pair, so the seed may not be a very strong bond, but in its predicted MFE structure all but 6 bases bind with the target site, so the overall binding seems promising.

The third best scoring seed match, at position 4325, also seems very promising with a high ΔG_{duplex} value of -23.83 and $\Delta\Delta G$ value of -9.61. In its predicted MFE structure, it has 14 base pairs in a row, which would form a very strong bond.

I would recommend further investigation into these three target sites to determine whether they are functional binding locations for svRNAa (position 5005) and svRNAb (position 4771 and 4325). The fourth best scoring seed match (position 4997) overlaps with the seed match at position 5005 and has a much less $\Delta\Delta G$ and ΔG_{duplex} value, so it unlikely to be functional binding site. The rest of the predicted seed matches have a high $\Delta\Delta G$ value and are unlikely to be functional binding sites.

PTGS-2

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
PTGS-2 3' UTR	svRNAa	2100	6:0:1	-17.4	-7.6	-9.79
PTGS-2 3' UTR	svRNAa	2002	7:0:1	-16.6	-10.92	-5.67
PTGS-2 3' UTR	svRNAb	239	6:0:1	-10.5	-10.9	0.4

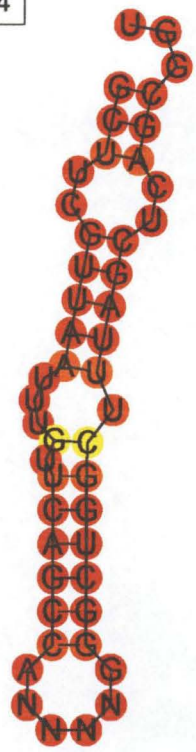
PTGS-2 Position 4052 Seed Match (3' UTR Position 2100)

One-to-One Alignment

```
                    4052
                     |
5' GCUUCGUUAAUUUGUUCAGCCA  3'  <- PTGS-2
   :|: |      |  : :|||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      4052
                       |
5'    GCUUCGUUAAUUUGUUCAGCCA  3'  <- PTGS-2
      |||  |:|||   | :||||||
3' UGGCGACUCGAUUU--C-GGUCGGG  5'  <- svRNAa
```



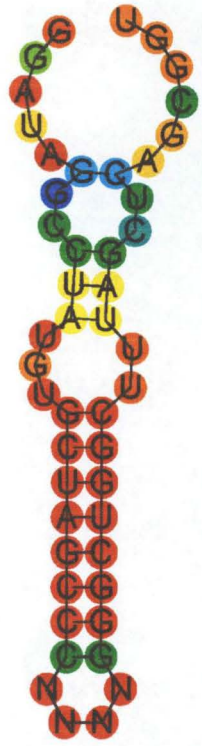
PTGS-2 Position 3954 Seed Match (3' UTR Position 2002)

One-to-One Alignment

```
                   3954
                    |
5' GGAUAGGCCUAUGUGCUAGCCC  3'  <- PTGS-2
   :     |  :  : ||:||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                    3954
                     |
5'  GGAUAGGCCUAUGUGCUAGCCC  3'  <- PTGS-2
        |  |||   ||:||||||
3' UGGCGACUCGAUUU-CGGUCGGG  5'  <- svRNAa
```



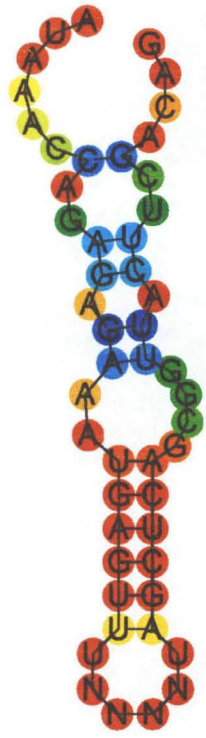
PTGS-2 Position 2191 Seed Match (3' UTR Position 239)

One-to-One Alignment

```
                      2191
                       |
5' AUAAACCAGAGAGAAAUGAGUUU  3'  <- PTGS-2
    |      || :|     ||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                       2191
                        |
5' AUAAACCAGAGAGA--AAUGAGUUU  3'  <- PTGS-2
         |   || :|     ||||:|
3'   GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



PTGS-2 Result Remarks

The best scoring seed match at position 4052 seems like a probable candidate for functional binding. Although it is a 6-mer seed match with a G:U wobble pair, it has low ΔG_{duplex} and $\Delta\Delta G$ value, and all but 6 of the svRNAa bases pair with the target site in the predicted MFE configuration. All of the base pairs have a very high base pair probability as well, so this seed match is a good candidate for a functional binding and should be further investigated and tested.

The other two predicted seed matches have high $\Delta\Delta G$ values and are most likely not able to form functional binding. The base pair probabilities of the base pairs outside the seed region in the MFE configurations for both matches are fairly low, and both matches only have around 50% of the bases in the svRNA pairing to the target site. Therefore, these matches are probably not functional binding sites.

TEP1

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
TEP1 3' UTR	svRNAa	1306	6:0:1	-15.41	-8.57	-6.83
TEP1 3' UTR	svRNAa	1350	6:0:1	-20.4	-14.6	-5.79
TEP1 3' UTR	svRNAa	1216	6:0:1	-18.19	-12.71	-5.47
TEP1 3' UTR	svRNAa	2177	6:0:1	-12.51	-7.87	-4.63
TEP1 3' UTR	svRNAa	1449	6:0:0	-15.92	-12.99	-2.92
TEP1 3' UTR	svRNAa	694	6:0:1	-18.9	-16.73	-2.16
TEP1 3' UTR	svRNAb	1770	7:0:1	-14.2	-12.57	-1.62
TEP1 3' UTR	svRNAa	776	6:0:0	-19.34	-18.32	-1.01
TEP1 3' UTR	svRNAb	1738	6:0:1	-10.36	-9.67	-0.68
TEP1 3' UTR	svRNAb	2471	6:0:1	-13.1	-13.45	0.35

TEP1 Position 9230 Seed Match (3' UTR Position 1306)

One-to-One Alignment

```
                    9230
                     |
5' GGUUAUCUGCACACCUCAGCCU  3'  <- TEP1
   : :  |  ||  |  :|||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                    9230
                     |
5'   GGUUAUCUGCACACCUCAGCCU  3'  <- TEP1
      |:|    ||  |  :|||||:
3' UGGCGACU--CGAUUUCGGUCGGG  5'  <- svRNAa
```

TEP1 Position 9274 Seed Match (3' UTR Position 1350)

One-to-One Alignment

```
                     9274
                      |
5' CGUGAGCCAUUGUGCCCGGCCC  3'  <- TEP1
     :|     :|: : ||:||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       9274
                        |
5'   CG-UGAGCCAUUGUGCCCGGCCC  3'  <- TEP1
     || |||||       | ||:||||
3' UGGCGACUCG--AUUUC-GGUCGGG  5'  <- svRNAa
```

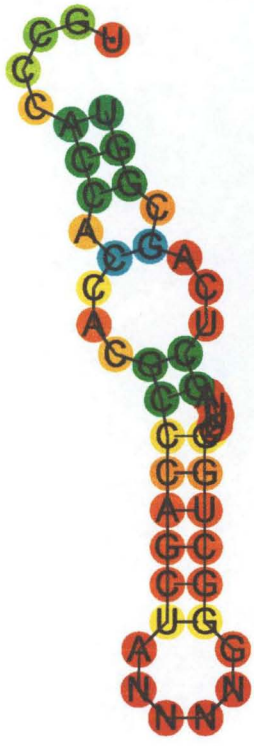
TEP1 Position 9140 Seed Match (3' UTR Position 1216)

One-to-One Alignment

```
                     9140
                      |
5' UGCCCACCACCACGCCCAGCUA  3'  <- TEP1
     | |    | | : |||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                        9140
                         |
5' UGCCCACCACCACGC-----CCAGCUA  3'  <- TEP1
        ||| |   ||     |||||:
3'      UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



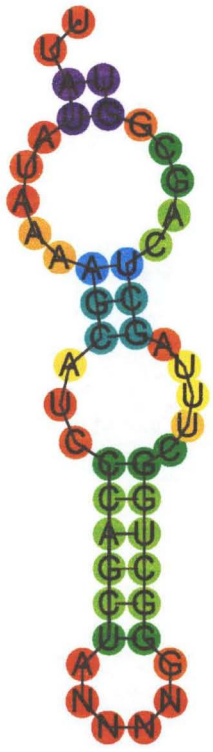
TEP1 Position 10101 Seed Match (3' UTR Position 2177)

One-to-One Alignment

```
                      10101
                      |
5'  UUAUAUAAAAGCAUCCCAGCUA  3'  <- TEP1
    :: : | |    |  |||||:
3' UGGCG-ACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      10101
                      |
5' UUAUAUAAAAGCAUC--CCAGCUA  3'  <- TEP1
             |||     |||||:
3'   UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



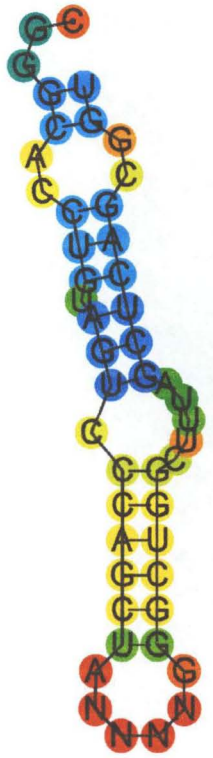
TEP1 Position 9373 Seed Match (3' UTR Position 1449)

One-to-One Alignment

```
                     9373
                     |
5' CGGGCACCUGUAGUCCCAGCUA  3'  <- TEP1
      ||     ||:  |||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      9373
                      |
5' CGGGCACCUGUAGU----CCCAGCUA  3'  <- TEP1
      :|  ||| ||:     |||||:
3'    UGGCGAC-UCGAUUUCGGUCGGG  5'  <- svRNAa
```



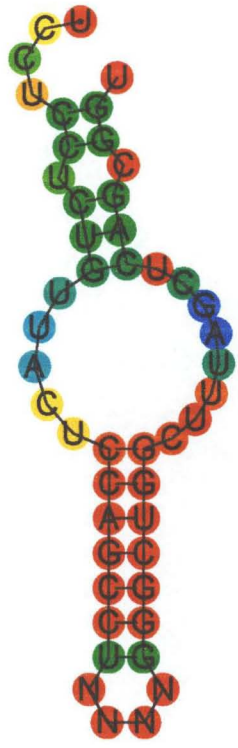
TEP1 Position 8618 Seed Match (3' UTR Position 694)

One-to-One Alignment

```
                    8618
                     |
5' UCCUCCUCUGUUACUCCAGCCU  3'  <- TEP1
    || |     | |  ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      8618
                       |
5' UCCUCCUCUGUUACU---CCAGCCU  3'  <- TEP1
       || |||         ||||||:
3'    UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



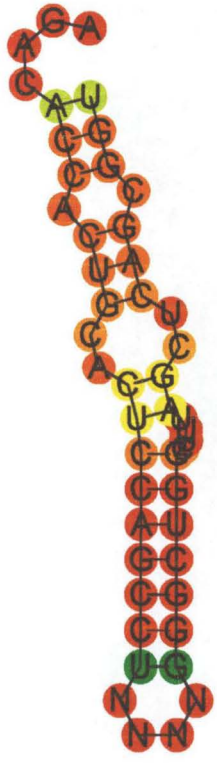
TEP1 Position 8700 Seed Match (3' UTR Position 776)

One-to-One Alignment

```
                    8700
                     |
5' AGACACCACUGCACUCCAGCCU  3'  <- TEP1
   |      | :  |  ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      8700
                       |
5' AGACACCACUGCACU----CCAGCCU  3'  <- TEP1
       ||| |||  ||    ||||||:
3'     UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



TEP1 Position 9694 Seed Match (3' UTR Position 1770)

One-to-One Alignment

```
                     9694
                      |
5' GGGAGUUAAAGGGCACUGAGUUU  3'  <- TEP1
     |    |  ::  | |||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      9694
                       |
5' GGGAGUUAAAG-GGC-ACUGAGUUU  3'  <- TEP1
       ||   ||| ::|   |||||:|
3'   GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

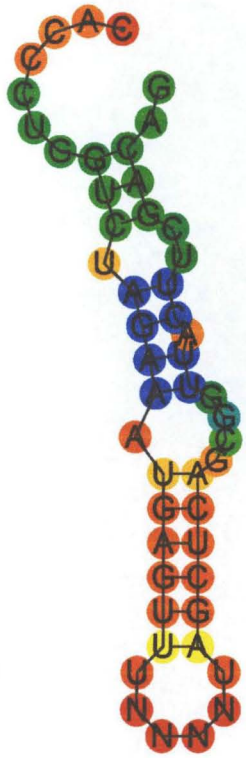
TEP1 Position 9662 Seed Match (3' UTR Position 1738)

One-to-One Alignment

```
                     9662
                      |
5' CACCCUGGUCUAGAAAUGAGUUU  3'  <- TEP1
   |    | ::    |    ||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                         9662
                          |
5' CACCCUGGUCUAG-AAA---UGAGUUU  3'  <- TEP1
       |||    || ||    ||||:|
3'     GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



TEP1 Position 10395 Seed Match (3' UTR Position 2471)

One-to-One Alignment

```
                      10395
                      |
5' UCUCCCAAAGGCUGGAUGAGUUU  3'  <- TEP1
   :   | ||  : : | ||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      10395
                      |
5'  UCUCCCAAAGGCUGGAUGAGUUU  3'  <- TEP1
    | ||  |  ::|:|  ||||:|
3' GACAGCUUCAUUGGCG-ACUCGAU  5'  <- svRNAb
```

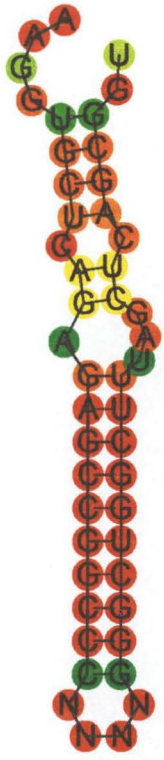
TEP1 Result Remarks

None of the TEP1 predicted target sites look promising for functional binding sites. 70% of the seed matches are 6-mer matches with a G:U wobble pair, which is not a very good seed match. None of the matches have a ΔΔG value below -10, and most of the matches are above a -5 ΔΔG value. While the TEP1 minor vault protein is required for a stable association between the vault RNA and the vault particle, and may help recruit the vault RNA into the vault particle, it does not appear that the small vault RNAs regulate the TEP1 gene.

NFKB1

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
NFKB1 3' UTR	svRNAa	99	8:0:1	-24.96	-13.38	-11.57
NFKB1 3' UTR	svRNAa	266	7:0:1	-21.2	-14.69	-6.5

NFKB1 Position 3476 Seed Match (3' UTR Position 99)



One-to-One Alignment

```
                  3476
                   |
5' AAGGUGCUCAGAGAGCCGGCCC   3'   <- NFKB1
   |   |:        |:||||:||||
3' UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                   3476
                    |
5' AAGGUGCUCAG--AGAGCCGGCCC   3'   <- NFKB1
      :|||  ||    :||||:||||
3'    UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```

NFKB1 Position 3643 Seed Match (3' UTR Position 266)

One-to-One Alignment

```
                  3643
                   |
5' CUUACUAAGCUUUUGCCAGCUG   3'   <- NFKB1
    :   ||  ||||    ||||||:
3' UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                  3643
                   |
5' CUUACUAAGCUUUUGCCAGCUG   3'   <- NFKB1
    :   ||  ||||    ||||||:
3' UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```

NFKB Result Remarks

Both predicted seed matches for the NFKB1 gene look very promising, especially the first ranking seed match. The first ranking seed match at position 3476 has an 8-mer seed match with one G:U wobble pair, and in its predicted MFE binding configuration there is 10 bases pairs in a row around the seed region, which would form a very strong bond. The position 3476 seed match has a very high ΔG_{duplex} value of -24.96 and a $\Delta\Delta G$ value of -11.57. All but six of the bases in svRNAa bind to the target site on the NFKB mRNA in the MFE binding configuration, and the bases have good base pair probabilities. This target site is definitely a candidate for functional binding, and could be further investigated and tested to see if svRNAa will bind to the target site and regulate the NFKB1 gene.

The second target site, at position 3643, while not as promising as the first, still has potential to be a functional binding site for svRNAa. It does not have as low of a $\Delta\Delta G$ value (-6.5), but the ΔG_{duplex} value is good, at -21.2, and it is a 7-mer seed match with one G:U wobble pair, which is a good seed match. The predicted MFE binding configuration is a straight one-to-one base pairing of svRNAa on the target site, and base pair probabilities in the binding are very high. For these reasons, I believe the target site at position 3643 is a candidate for functional binding, even though the $\Delta\Delta G$ value is not as low as the recommended -10 value for functional binding.

NFKB2

No svRNAa or svRNAb seed matches were found on the NFKB2 mRNA.

RELA

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
RELA 3' UTR	svRNAa	433	6:0:0	-14.9	-4.74	-10.15
RELA 3' UTR	svRNAa	308	6:0:1	-15.22	-18.5	3.28
RELA 3' UTR	svRNAa	687	6:0:1	-19	-22.44	3.44

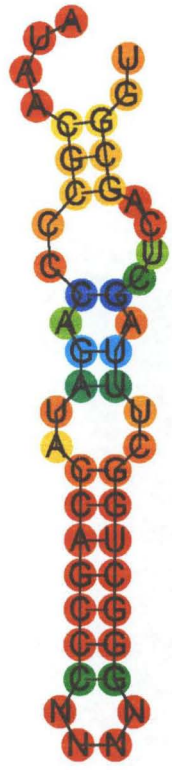
RELA Position 2229 Seed Match (3' UTR Position 433)

One-to-One Alignment

```
                    2229
                     |
5' AUAACGCCCCAGAUACCAGCCC  3'  <- RELA
   |:  |    | :|  |||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                     2229
                      |
5' AUAACGCCC--CAGAUACCAGCCC  3'  <- RELA
       |||    | :|  |||||||
3'   UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



RELA Position 2104 Seed Match (3' UTR Position 308)

One-to-One Alignment

```
                    2104
                     |
5' CCCCAUCCCCAUCCUCCAGCUU  3'  <- RELA
    ||  |   |      |||||::
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                    2104
                     |
5' CCCCAUCCCCAUCCUCCAGCUU  3'  <- RELA
                   |||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

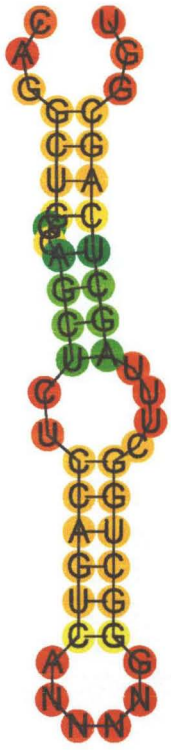
RELA Position 2483 Seed Match (3' UTR Position 687)

One-to-One Alignment

```
                    2104
                     |
5' CAGGCUGGCAGCUCUCCAGUCA  3'  <- RELA
     ||||:        ||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                     2104
                      |
5' CAGGCUGGCAGCUCU--CCAGUCA  3'  <- RELA
     ||||  ||||    ||||||
3' UGGCGAC--UCGAUUUCGGUCGGG  5'  <- svRNAa
```



RELA Result Remarks

The predicted target sites for the RELA mRNA are not very promising for functional binding sites for svRNA. The top ranking target site, at position 2229, does have some potential for being functional binding site. The seed match is a 6-mer perfect match, but the ΔG_{duplex} value for the target site is not very low, at only -14.9. The $\Delta\Delta G$ is low though, at -10.15, since the ΔG_{duplex} value was relativity high, ΔG_{open} had to be high as well (which it was, at a value of -4.74). The base pair probabilities in the predicted MFE binding configuration are not very high, so it is unclear whether this target site could be a functional site or not.

The other two predicted target sites are not good candidates for functional binding. They had very low ΔG_{open} values, causing them to have high $\Delta\Delta G$ values. It is unlikely that svRNAa would be able to bind to these target sites.

RELB

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
RELB 3' UTR	svRNAa	104	6:0:1	-15.7	-10.61	-5.08
RELB 3' UTR	svRNAb	369	6:0:1	-9.3	-12.14	2.84

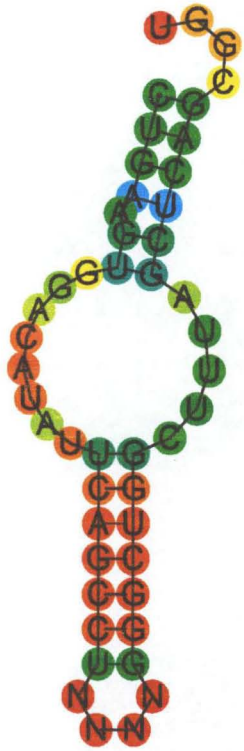
RELB Position 1970 Seed Match (3' UTR Position 104)

One-to-One Alignment

```
                    1970
                     |
5' CUGAAGUGGACAUAUUCAGCCU  3'  <- RELB
    :     :|  | | :|||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                         1970
                          |
5'      CUGAAGUGGACAUAUUCAGCCU  3'  <- RELB
       |||| |:         :||||||:
3' UGGCGACU-CGAUUUC---GGUCGGG  5'  <- svRNAa
```



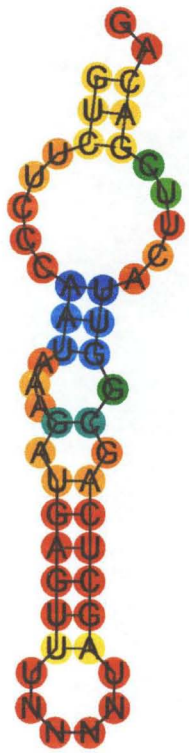
RELB Position 2235 Seed Match (3' UTR Position 369)

One-to-One Alignment

```
                    2235
                     |
5' GUCUUCCCAAUAAAGAUGAGUUU  3'  <- RELB
    | |:       |  | ||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      2235
                       |
5'   GUCUUCCCAAUAAAGAUGAGUUU  3'  <- RELB
     |||      ||:   | ||||:|
3' GACAGCUUCAUUGG--CGACUCGAU  5'  <- svRNAb
```



RELB Result Remarks

The predicted target sites for RELB do not have potential to be functional binding sites for the svRNAs. Their ΔΔG value is too high, and the base pair probabilities of the bases in the predicted MFE binding configuration are low.

REL

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	ΔΔG
REL 3' UTR	svRNAa	477	6:0:0	-21.83	-12.04	-9.78
REL 3' UTR	svRNAb	452	6:0:0	-16.9	-12.27	-4.62
REL 3' UTR	svRNAa	311	6:0:0	-18.11	-15.49	-2.61

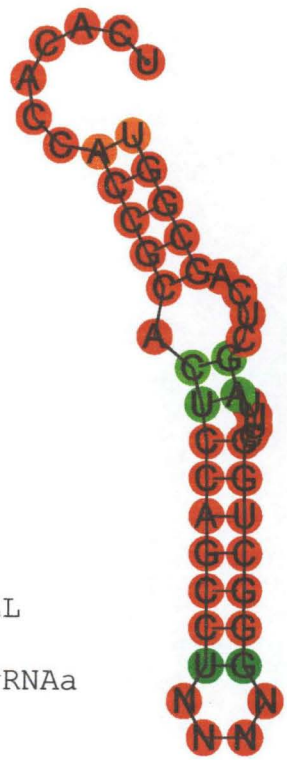
REL Position 2561 Seed Match (3' UTR Position 477)

One-to-One Alignment

```
                    2561
                     |
5'  UCACACCACCGCACUCCAGCCU   3'   <- REL
     |      | | :|  ||||||:
3'  UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                           2561
                            |
5'  UCACACCACCGC---ACU----CCAGCCU   3'   <- REL
          |||||    ||    ||||||:
3'        UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```



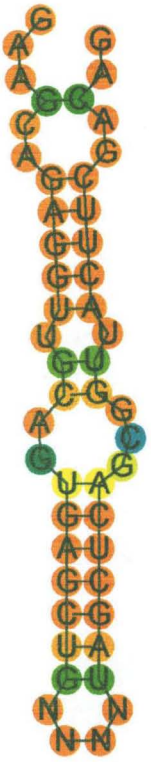
REL Position 2536 Seed Match (3' UTR Position 452)

One-to-One Alignment

```
                      2536
                       |
5' GAAGCAGAGGUUGCAGUGAGCUG  3'  <- REL
       | :||    |  ||||||:
3' GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      2536
                       |
5' GAAGCAGAGGUUGC-AGUGAGCUG  3'  <- REL
      |  ||:|| :|   ||||||:
3'  GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```



REL Position 2395 Seed Match (3' UTR Position 311)

One-to-One Alignment

```
                     2395
                      |
5' ACCAGGAAUUCGAGACCAGCCU  3'  <- REL
   |||    | : :|: ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                     2395
                      |
5'  ACCAG-GAAUUCGAGACCAGCCU  3'  <- REL
    |||   || :| :|| ||||||:
3'  UGGCGACUCGAUUUC-GGUCGGG  5'  <- svRNAa
```

REL Result Remarks

The top ranking target site for REL is a good candidate for functional binding of svRNAa to the target site at position 2561 of the REL mRNA. It has a low ΔG_{duplex} value of -21.83 and a low $\Delta\Delta G$ value of -9.78. The seed match is a perfect 6-mer match, and the base pair probabilities of the bases in the predicted MFE binding structure are high.

The second best ranking target site for REL may also be a candidate for functional binding, except with svRNAb instead of svRNAa. While the $\Delta\Delta G$ value is somewhat high, at -4.62, the base pair probabilities of the predicted MFE binding configuration are quite high, and the MFE configuration has a 7 base pair (6-mer seed match plus 1 G:U wobble pair) and a 5 base pair sequence that make for a nice binding. Interesting, the target site for svRNAb (seed position 2536) on the REL mRNA is only 10 nucleotides upstream from where svRNAa binds to the predicted target site (seed position 2561).

This makes me wonder if both these target sites are functional bindings sites, the first for svRNAa and the second for svRNAb, making this region of the 3' UTR a svRNA regulatory region. Further investigation and testing could be done to verify whether these target sites are functional binding sites for svRNAa and svRNAb.

NLRP1

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
NLRP1 3' UTR	svRNAa	181	6:0:1	-18.51	-17.23	-1.27

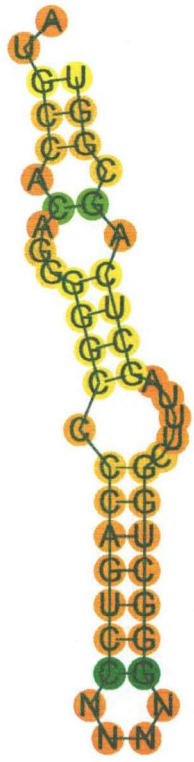
NLRP1 Position 5158 Seed Match (3' UTR Position 181)

One-to-One Alignment

```
                     5158
                     |
5' AUGCCACAGGGGGCCCCAGUCC  3'  <- NLRP1
   |:  |  ||  ::  ||||:||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       5158
                       |
5' AUGCCACAGGGGGCC----CCAGUCC  3'  <- NLRP1
     :|| |   |:||     ||||:||
3'   UGGCG--ACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



NLRP1 Result Remarks

The predicted target site on the NLRP1 mRNA is most likely not a functional binding site for svRNAa. The free energy lost when unpairing the target site from the mRNA secondary structure is almost as much as the free energy gained from binding svRNAa to the target site, so this binding will probably not happen. Also, the seed match is 6-mer match with one G:U wobble pair, which is not a very strong bond, and the base pair probabilities for the MFE binding configuration are not extremely high. For these reason, I doubt this target site is a functional binding site for svRNAa.

PRKACB

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
PRKACB 3' UTR	svRNAa	2216	6:0:1	-19.7	-8.72	-10.97
PRKACB 3' UTR	svRNAa	1092	6:0:1	-18.6	-7.66	-10.93
PRKACB 3' UTR	svRNAa	698	6:0:0	-21.6	-10.86	-10.73
PRKACB 3' UTR	svRNAa	2558	6:0:1	-15.7	-6.79	-8.9
PRKACB 3' UTR	svRNAa	2155	6:0:1	-12.9	-6.37	-6.52
PRKACB 3' UTR	svRNAa	2351	6:0:1	-10.8	-5.16	-5.63
PRKACB 3' UTR	svRNAb	19	7:0:0	-15.11	-10.77	-4.33

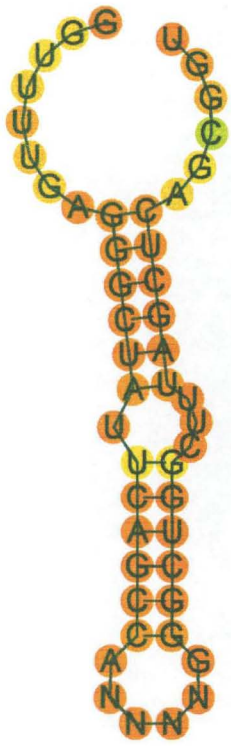
PRKACB Position 3506 Seed Match (3' UTR Position 2216)

One-to-One Alignment

```
                    3506
                     |
5' GGUUUUGAGGGCUAUUCAGCCA  3'  <- PRKACB
   : : :||||     | :||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                     3506
                      |
5' GGUUUUGAGGGCUA--UUCAGCCA  3'  <- PRKACB
          |:|||||   :||||||
3'   UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



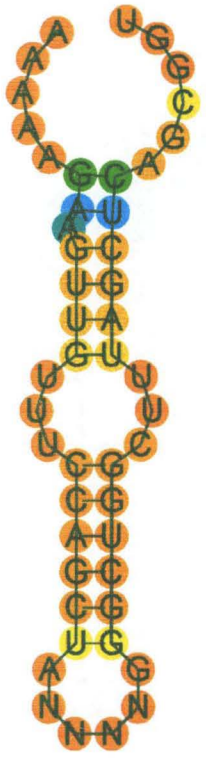
PRKACB Position 2382 Seed Match (3' UTR Position 1092)

One-to-One Alignment

```
                    2382
                     |
5' AAAAAGAAGUUGUUUCCAGCUA  3'  <- PRKACB
   |       ||:|:   |||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                    2382
                     |
5'  AAAAAGAAGUUGUUUCCAGCUA  3'  <- PRKACB
        || |:|:   |||||:
3' UGGCGACU-CGAUUUCGGUCGGG  5'  <- svRNAa
```



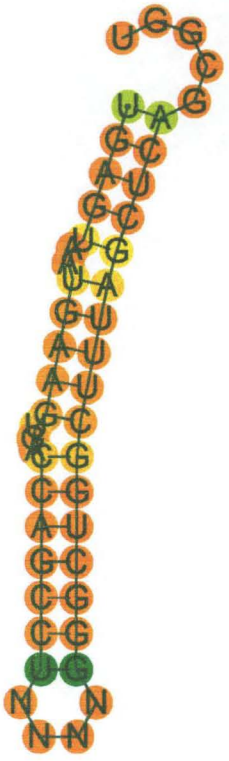
PRKACB Position 1988 Seed Match (3' UTR Position 698)

One-to-One Alignment

```
                    1988
                     |
5' UGAGUAAUGAAGUGACCAGCCU  3'  <- PRKACB
     |:   |  : : ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      1988
                       |
5'      UGAGUAAUGAAGUGACCAGCCU  3'  <- PRKACB
        ||||:  |:|||   ||||||:
3' UGGCGACUCG--AUUUC---GGUCGGG  5'  <- svRNAa
```



PRKACB Position 3848 Seed Match (3' UTR Position 2558)

One-to-One Alignment

```
                      3848
                       |
5' ACUAUCAGUAUGAUCUCAGCCU  3'  <- PRKACB
   ||: :  :  |:|  :||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      3848
                       |
5' ACUA-UCAGUAUGAUCUCAGCCU  3'  <- PRKACB
   ||:  | ||: |:|  :||||||:
3' UGGCGACUCG-AUUUCGGUCGGG  5'  <- svRNAa
```

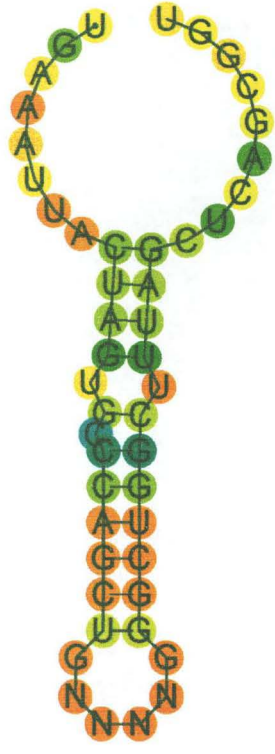
PRKACB Position 3445 Seed Match (3' UTR Position 2155)

One-to-One Alignment

```
                      3445
                       |
5' UGAAAUUACUAGUGCCCAGCUG  3'  <- PRKACB
       | | : : : ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      3445
                       |
5'  UGAAAUUACUAGUGCCCAGCUG  3'  <- PRKACB
            |||: | ||||||:
3' UGGCGACUCGAUUUC-GGUCGGG  5'  <- svRNAa
```



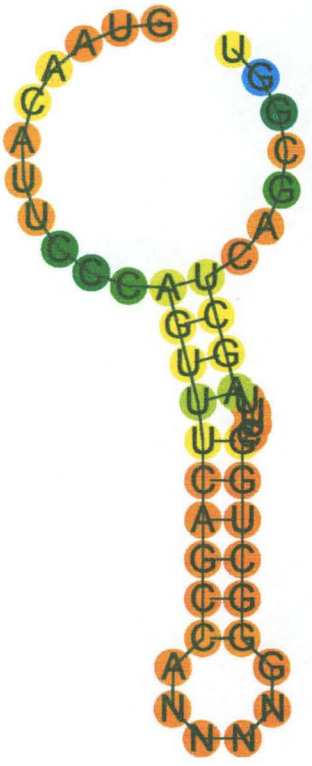
PRKACB Position 3641 Seed Match (3' UTR Position 2351)

One-to-One Alignment

```
                    3641
                    |
5' GUAACAUUCCCAGUUUCAGCCA  3'  <- PRKACB
   ::   |     | |:  :||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      3641
                      |
5' GUAACAUUCCCAGUU----UCAGCCA  3'  <- PRKACB
              ||:|    :|||||
3'     UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



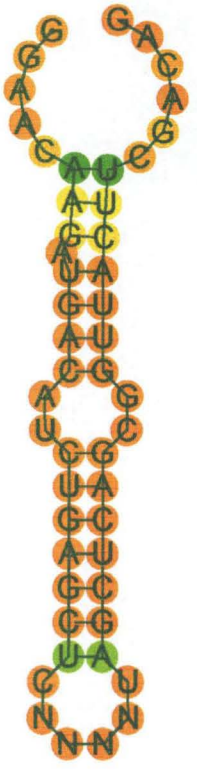
PRKACB Position 1309 Seed Match (3' UTR Position 19)

One-to-One Alignment

```
                    1309
                    |
5' GGAACAAGAUGACAUCUGAGCUC  3'  <- PRKACB
       | |: |:||  |||||||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                     1309
                     |
5'  GGAACAAGAUGACAUCUGAGCUC  3'  <- PRKACB
         ||| |:||  |||||||
3' GACAGCUUC-AUUGGCGACUCGAU  5'  <- svRNAb
```



PRKACB Result Remarks

The top three ranking predicted target sites for PRKACB seem like promising candidates for functional binding sites for svRNAa. All three have a ΔΔG value less than -10, and low ΔG_{duplex} values. The top two target sites have 6-mer seed matches with one G:U wobble pair though. But the base pair probabilities of the bases in the predicted MFE binding configurations are high for all three target sites. These factors indicate that these target sites have potential to be a functional binding site for svRNAa, which could regulate the PRKACB gene. Further investigation and testing could be conducted to determine whether these predicted target sites are functional binding sites for svRNAa.

The fourth ranking target site has a ΔΔG value of -8.9, which is near the recommended -10 ΔΔG value for functional binding sites. But the base pair probabilities for the bases in the MFE predicted binding configuration are very low, indicating that the target site is probably not a functional binding site. The rest of the predicted target sites are most likely nonfunctional, as their ΔΔG and ΔG_{duplex} values are high.

IL10 (Interleukin 10)

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	ΔΔG
IL10 3' UTR	svRNAa	569	7:0:1	-22	-9.54	-12.45
IL10 3' UTR	svRNAb	385	8:0:1	-20.18	-9.24	-10.93
IL10 3' UTR	svRNAa	691	6:0:0	-20.05	-11.55	-8.49
IL10 3' UTR	svRNAb	427	7:0:1	-14.05	-6.44	-7.6
IL10 3' UTR	svRNAa	84	6:0:0	-20.1	-16.08	-4.01
IL10 3' UTR	svRNAb	168	7:0:0	-11.1	-7.53	-3.56
IL10 3' UTR	svRNAa	772	6:0:1	-16.82	-13.45	-3.36
IL10 3' UTR	svRNAa	857	6:0:0	-21.81	-21.4	-0.4
IL10 3' UTR	svRNAb	832	6:0:0	-16.6	-19.64	3.04

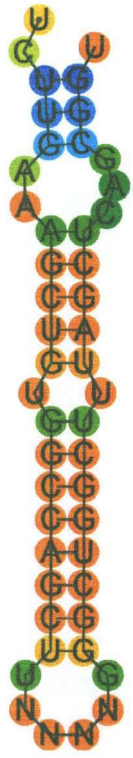
IL10 Position 1165 Seed Match (3' UTR Position 569)

One-to-One Alignment

```
                    1165
                     |
5' UCUUGAAAGCUGUGGCCAGCUU  3'  <- IL10
    |:    ||||: :|||||||::
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                    1165
                     |
5' UCUUGAA-AGCUGUGGCCAGCUU  3'  <- IL10
     ::|   ||||: :|||||||::
3'  UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



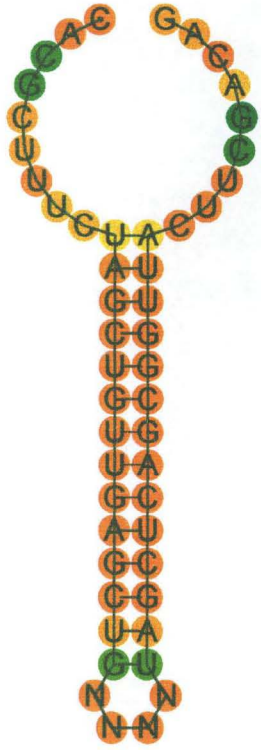
IL10 Position 981 Seed Match (3' UTR Position 385)

One-to-One Alignment

```
                    981
                     |
5' CACGCUUUCUAGCUGUUGAGCUG  3'  <- IL10
   |   |    ||:|:|:|||||||:
3' GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

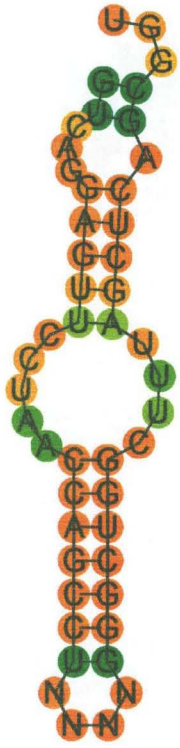
```
                    981
                     |
5' CACGCUUUCUAGCUGUUGAGCUG  3'  <- IL10
            ||:|:|:|||||||:
3' GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```



IL10 Position 1287 Seed Match (3' UTR Position 691)

One-to-One Alignment

```
                    1287
                     |
5' GUCAGGAGUUCCUAACCAGCCU  3'  <- IL10
   ::|    : :    | ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



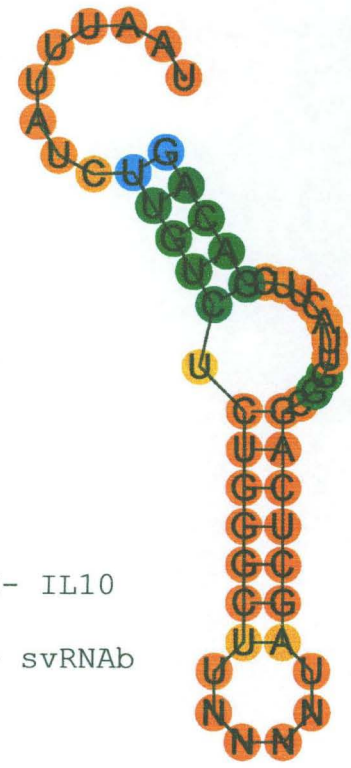
RNAfold MFE Predicted Alignment

```
                      1287
                       |
5'    GUCAGGAGUUCCUAACCAGCCU  3'  <- IL10
      |:   |||:|     ||||||:
3' UGGCG--ACUCGAUUUC-GGUCGGG  5'  <- svRNAa
```

IL10 Position 1023 Seed Match (3' UTR Position 427)

One-to-One Alignment

```
                    1023
                     |
5' UAAUUUAUCUUGUCUCUGGGCUU  3'  <- IL10
   :  |: |  | ::| |||:|||
3' GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```



RNAfold MFE Predicted Alignment

```
                           1023
                            |
5' UAAUUUAUCUUGUC---------UCUGGGCUU  3'  <- IL10
            :||||          |||:|||
3'          GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```

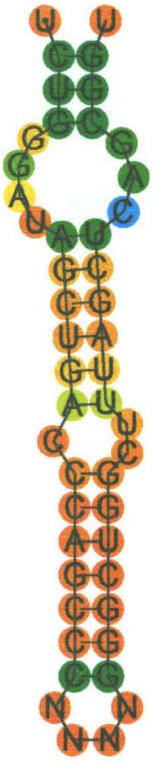
IL10 Position 680 Seed Match (3' UTR Position 84)

One-to-One Alignment

```
                    680
                     |
5' UCUGGGAUAGCUGACCCAGCCC  3'  <- IL10
    |:|         :| |||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                    680
                     |
5' UCUGGGAUAGCUGAC-CCAGCCC  3'  <- IL10
    |:|    ||||:|  |||||||
3' UGGCGAC-UCGAUUUCGGUCGGG  5'  <- svRNAa
```



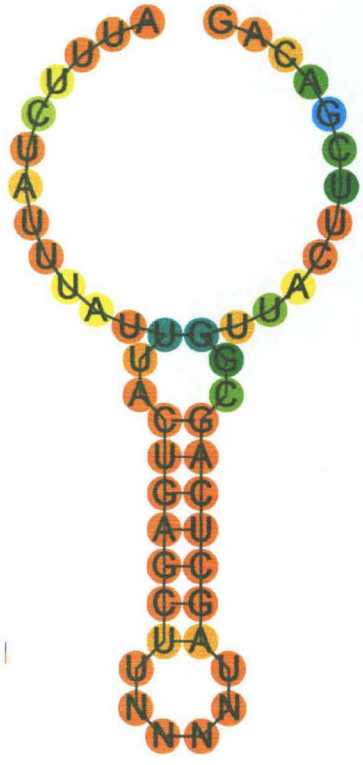
IL10 Position 764 Seed Match (3' UTR Position 168)

One-to-One Alignment

```
                    764
                     |
5' AUUUCUAUUUAUUUACUGAGCUU  3'  <- IL10
    | || |  || :: |||||||
3' GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                    764
                     |
5' AUUUCUAUUUAUUUACUGAGCUU  3'  <- IL10
                 :  |||||||
3' GACAGCUUCAUUGGCGACUCGAU 5'  <- svRNAb
```



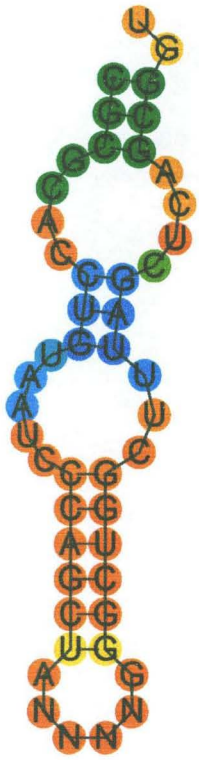
IL10 Position 1368 Seed Match (3' UTR Position 772)

One-to-One Alignment

```
                     1368
                     |
5' CGCGCACCUGUAAUCCCAGCUA  3'  <- IL10
     |||      |||  |||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                      1368
                      |
5'   CGCGCACCUGUAAUCCCAGCUA  3'  <- IL10
     |||     ||:     |||||:
3' UGGCGACUCGAUUUC--GGUCGGG  5'  <- svRNAa
```



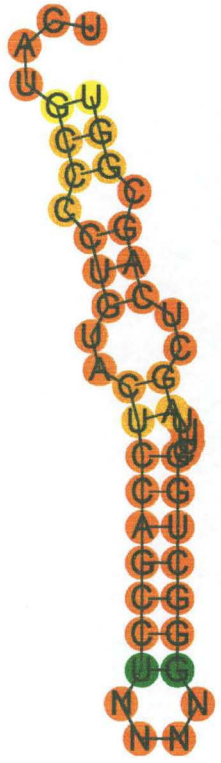
IL10 Position 1453 Seed Match (3' UTR Position 857)

One-to-One Alignment

```
                     1453
                     |
5' UCAUGCCCCUGUACUCCAGCCU  3'  <- IL10
    |        : ||  ||||||:
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       1453
                       |
5' UCAUGCCCCUGUACU----CCAGCCU  3'  <- IL10
       :|| |||  ||    ||||||:
3'     UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



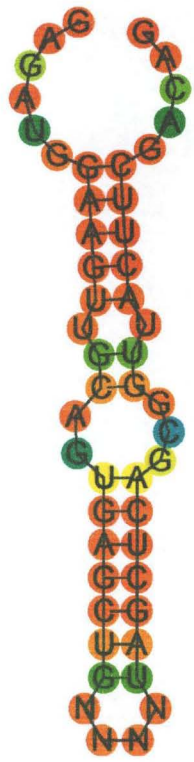
IL10 Position 1428 Seed Match (3' UTR Position 832)

One-to-One Alignment

```
                      1428
                      |
5' GAGAUGGAAGUUGCAGUGAGCUG  3'  <- IL10
     | ||:|     |  ||||||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      1428
                      |
5' GAGAUGGAAGUUGC-AGUGAGCUG  3'  <- IL10
         |||||  :|   ||||||:
3'  GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



IL10 Result Remarks

The top three ranking target sites for IL10 seem like good candidates for functional binding sites for svRNAs. The top ranking target site, at position 1165, has a 7-mer seed match with one G:U wobble, with a good ΔΔG value of -12.45 and a good ΔG_{duplex} value of -22. All but 5 of the bases in svRNAa bind to the target site in the predicted MFE binding configuration. The second ranking target site has an 8-mer seed with one G:U wobble, and good values of 10.93 for ΔΔG and -20.18 for ΔG_{duplex}. The predicted MFE binding configuration has 14 base pairs in a row on the 5' end of svRNAb, forming a very nice bond to the target site. The third ranking target site has a perfect 6-mer seed match, with a good ΔG_{duplex} value of -20.05, and a decent ΔΔG value of -8.49, slightly below the recommended -10 value. These three target sites could be further investigated and tested to determine whether they are functional binding sites from svRNAs and svRNAb.

The rest of predicted target sites from IL10 have little likelihood of being functional binding sites due to high ΔΔG values and low base pair probabilities in the predicted MFE binding configurations.

CXCL14

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
CXCL14 3' UTR	svRNAb	343	7:0:0	-14.4	-5.36	-9.03
CXCL14 3' UTR	svRNAb	373	6:0:0	-15.6	-12.36	-3.23
CXCL14 3' UTR	svRNAb	475	7:0:1	-13.5	-16.01	2.51

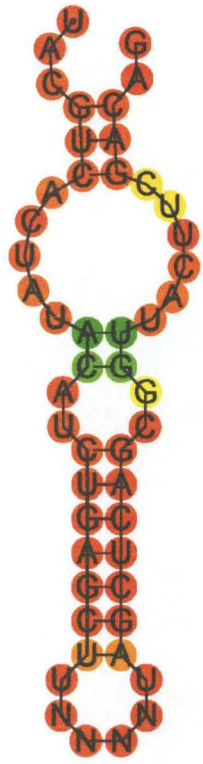
CXCL14 Position 1144 Seed Match (3' UTR Position 343)

One-to-One Alignment

```
                    1144
                     |
5' UACGUCACUAUACAUCUGAGCUU  3'  <- CXCL14
   :   : |    ||  |||||||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                    1144
                     |
5' UACGUCACUAU-ACAUCUGAGCUU  3'  <- CXCL14
      |||       ||  |||||||
3'  GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



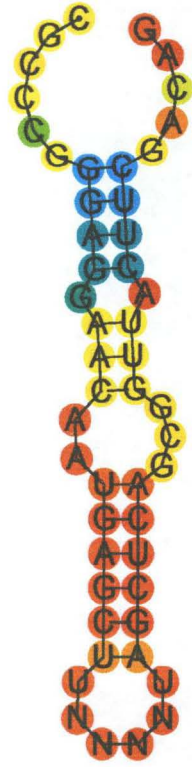
CXCL14 Position 1174 Seed Match (3' UTR Position 104)

One-to-One Alignment

```
                    1174
                     |
5' CGCCCGGGAGGAACAAUGAGCUU  3'  <- CXCL14
   |   ||::  :| |  ||||||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                    1174
                     |
5' CGCCCGGGAGGAACAA-UGAGCUU  3'  <- CXCL14
         |:|| |||   ||||||
3'  GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```



CXCL14 Position 1276 Seed Match (3' UTR Position 475)

One-to-One Alignment

```
                    1276
                    |
5' CUGCCACGGGCUCUCCUGGGCUU  3'  <- CXCL14
   ||| |  :|   |: |||:|||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                    1276
                    |
5' CUGCCACGGGCUCUCCUGGGCUU  3'  <- CXCL14
   ||| |  :|   |: |||:|||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

CXCL14 Result Remarks

The top ranking target site for CXCL14 is a good candidate for a functional binding site for svRNAb, which could regulate the CXCL14 gene. This would make sense for svRNAb to regulate CXCL14, since it plays a role in inflammation and in known to be found in breast cells. Persson *et al.* found CXCL14 to be upregulated in MCF-7 breast cancer cells, and they also found svRNAb to regulate the CYP3A4 gene in these MCF-7 cancer cells, so it makes sense that svRNAb would also regulate CXCL14 since it is so closely linked with breast cells. The target site on the CXCL14 mRNA has a 7-mer perfect seed match with svRNAb, and has a good ΔΔG value of -9.03, near the recommended -10 value for functional binding sites. The base pair probabilities for the predicted MFE binding configuration are high, and the structure looks well formed with 3 groups of binding base pairs and two unpaired loops. Further investigation and testing could be done to verify whether this target site in a functional binding site for svRNAb.

The other two predicted target sites have high ΔΔG values and low base pair probabilities for their predicted MFE binding configurations, and are therefore not likely functional binding sites.

ABCC4

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
ABCC4 3' UTR	svRNAb	1205	6:0:1	-18.11	-7.15	-10.95
ABCC4 3' UTR	svRNAb	1093	6:0:1	-12.89	-5.85	-7.03
ABCC4 3' UTR	svRNAa	158	6:0:1	-11.82	-5.43	-6.38
ABCC4 3' UTR	svRNAb	908	6:0:1	-15.01	-13.71	-1.29

ABCC4 Position 5302 Seed Match (3' UTR Position 1205)

One-to-One Alignment

```
                      5302
                       |
5' GGUGGCUGUAAUAUGUUGAGUUC  3'  <- ABCC4
          :   |   :|:|||||:|
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                      5302
                       |
5'  GGUGGCUGUAAUAUGUUGAGUUC  3'  <- ABCC4
     || |  |||||: :|:|||||:|
3' GACAGCUUCAUUG-GCGACUCGAU  5'  <- svRNAb
```

ABCC4 Position 5190 Seed Match (3' UTR Position 1093)

One-to-One Alignment

```
                      5190
                       |
5' UAAUAGCAUUGUCAUUUGGGCUU  3'  <- ABCC4
   :  | | | |: |  :||:|||
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                           5190
                            |
5' UAAUAGCAUUGUCAU--------UUGGGCUU  3'  <- ABCC4
          :||||           :||:|||
3'        GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

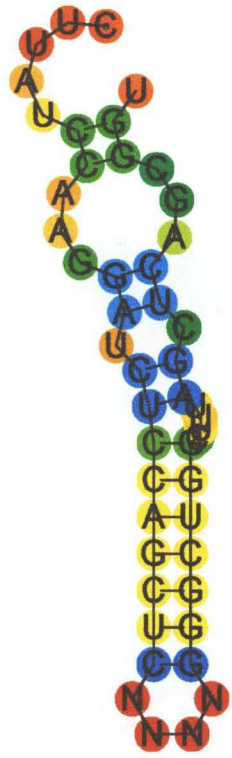
ABCC4 Position 4255 Seed Match (3' UTR Position 158)

One-to-One Alignment

```
                     4255
                     |
5' CUUAUCCAAGGAUCUCCAGCUC  3'  <- ABCC4
    :: :  |   |   |||||:|
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       4255
                       |
5' CUUAUCCAAGGAUCU----CCAGCUC  3'  <- ABCC4
       ||   || ||    |||||:|
3'     UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



ABCC4 Position 5005 Seed Match (3' UTR Position 908)

One-to-One Alignment

```
                    5005
                    |
5' UGUAAGCCUUUUGGUUUGGGCUG  3'  <- ABCC4
   :    |   |     :||:|||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                     5005
                     |
5' UGUAAGCCUUUUGGUUUGGGCUG  3'  <- ABCC4
                   :||:|||:
3' GACAGCUUCAUUGGCGACUCGAU  5'  <- svRNAb
```

ABCC4 Result Remarks

The top ranking target site on the ABCC4 mRNA, at position 5302, is a good candidate for functional binding of svRNAb. The seed match is only a 6-mer match with one G:U wobble pair, but the match has good values of -18.11 for ΔG_{duplex} and -10.95 for $\Delta\Delta G$. In the predicted MFE binding configuration, all but 6 of the svRNAb bases are paired with the target site, and 14 bases in a row are paired extending from the svRNAb seed. The base pair probabilities in MFE configuration are very high also. Further investigation and testing could be done to verify whether the target site at position 5302 of the ABCC4 mRNA is a functional binding site for svRNAb.

The other 3 predicted target sites are most likely not functional binding sites due to high $\Delta\Delta G$ values, and low base pair probabilities in the predicted MFE binding configurations.

ABCG2

Gene	microRNA	Position	Seed	ΔG_{duplex}	ΔG_{open}	$\Delta\Delta G$
ABCG2 3' UTR	svRNAa	1106	6:0:1	-14.4	-3.02	-11.37
ABCG2 3' UTR	svRNAa	1094	6:0:1	-12.1	-4.62	-7.47
ABCG2 3' UTR	svRNAb	1559	7:0:0	-16.2	-9.46	-6.73
ABCG2 3' UTR	svRNAa	1509	6:0:1	-9.83	-7.78	-2.04

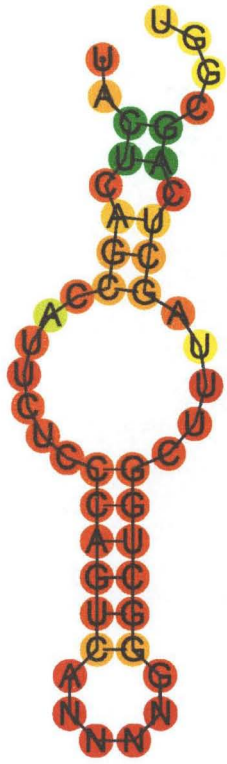
ABCG2 Position 3567 Seed Match (3' UTR Position 1106)

One-to-One Alignment

```
                      3567
                      |
5' UACUCAGCCAUUCUCCCAGUCA  3'  <- ABCG2
     |:| |    |     ||||:|
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       3567
                       |
5'   UACUCAGCCAUUCUCCCAGUCA  3'  <- ABCG2
       || |||         ||||:|
3' UGGCGACUCGAUUUC--GGUCGGG  5'  <- svRNAa
```



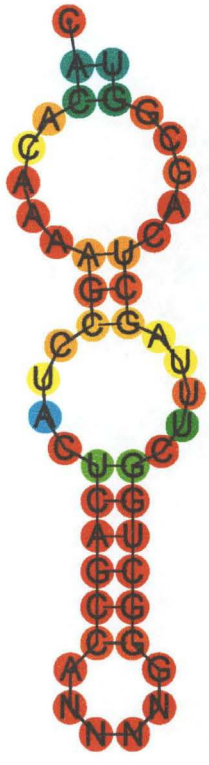
ABCG2 Position 3555 Seed Match (3' UTR Position 1094)

One-to-One Alignment

```
                      3555
                      |
5' CACACAAAAGCCUACUCAGCCA  3'  <- ABCG2
     | |  |      | :||||||
3' UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       3555
                       |
5' CACACAAAAGCCUAC-UCAGCCA  3'  <- ABCG2
    ||      |||      :|||||
3'  UGGCGACUCGAUUUCGGUCGGG  5'  <- svRNAa
```



ABCG2 Position 4020 Seed Match (3' UTR Position 1559)

One-to-One Alignment

```
                      4020
                      |
5' ACUUACAGGAGUUAACUGAGCUG   3'   <- ABCG2
       |   |:| : :   |||||||:
3' GACAGCUUCAUUGGCGACUCGAU   5'   <- svRNAb
```

RNAfold MFE Predicted Alignment

```
                       4020
                       |
5' ACUUACAGGAGUUAA--CUGAGCUG   3'   <- ABCG2
           |:|||      |||||||:
3'   GACAGCUUCAUUGGCGACUCGAU   5'   <- svRNAb
```

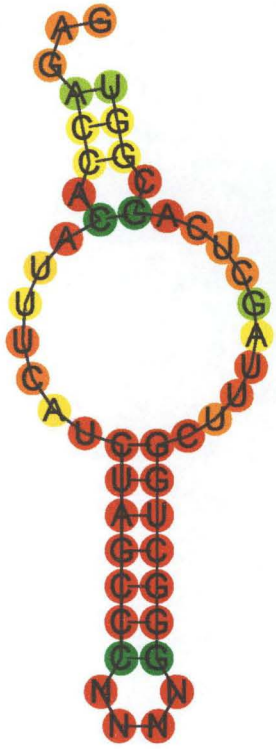
ABCG2 Position 3970 Seed Match (3' UTR Position 1509)

One-to-One Alignment

```
                      3970
                      |
5' GAGACCACAUUUCAUCUAGCCC   3'   <- ABCG2
   :    |     :|   | |:|||||
3' UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```

RNAfold MFE Predicted Alignment

```
                       3970
                       |
5' GAGACCACAUUUCAU---CUAGCCC   3'   <- ABCG2
      |||  |           |:|||||
3'    UGGCGACUCGAUUUCGGUCGGG   5'   <- svRNAa
```



ABCG2 Result Remarks

The top ranking target site of ABCG2, at position 3567, is a potential functional binding site from svRNAa, although it is not a very promising site. It has a relatively high ΔG_{duplex} value, and while it has a low $\Delta\Delta G$ value, this is only due to a very high ΔG_{open} value. The seed match for the target site is only a 6-mer match with a G:U wobble pair. The base pair probabilities of the predicted MFE binding configuration are not extremely high. So while the target site at position 3567 is a potential functional biding site, it is not the best candidate, and I recommend focusing on other potential target sites in other genes before further investigation of this target site.

The other target sites of ABCG2 show little potential for being a functional binding site for svRNAs due to high ΔG_{duplex} and $\Delta\Delta G$ values, and low base pair probabilities in the predicted MFE binding configurations.

Works Cited

Gruber, Andreas R., Ronny Lorenz, Stephan H. Bernhart, Richard Neuböck, and Ivo L. Hofacker. "The Vienna RNA Websuite." *Nucleic Acids Research* 36 (2008): 70-74. *PubMed Central.* National Center for Biotechnology Information, 1 July 2008. Web. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2447809/>.

Hofacker, Ivo L. "How MicroRNAs Choose Their Targets." *Nature Genetics* 39.10 (2007): 1191-192. Print.

Kertesz, Michael, Nicola Iovino, Ulrich Unnerstall, Ulrike Gaul, and Eran Segal. "The Role of Site Accessibility in MicroRNA Target Recognition." *Nature Genetics* 39.10 (2007): 1278-284. Print.

Kickhoefer, Valerie A., Kavitha S. Rajavel, George L. Scheffer, William S. Dalton, Rik J. Scheper, and Leonard H. Rome. "Vaults Are Up-regulated in Multidrug-resistant Cancer Cell Lines." *The Journal of Biological Chemistry* 273.April 10 (1998): 8971-974. Print.

Kickhoefer, Valerie A., Yie Liu, Lawrence H. Rome, Bryan E. Snow, Phoebe L. Stewart, Lea Harrington, and Leonard H. Rome. "The Telomerase/vault-associated Protein TEP1 Is Required for Vault RNA Stability and Its Association with the Vault Particle." *The Journal of Cell Biology* 152.1 (2001): 157-64. Print.

National Center for Biotechnology Information. National Library of Medicine, National Institutes of Health. Web. 2010. <http://www.ncbi.nlm.nih.gov/>.

"Online MicroRNA Prediction Tool." *Segal Lab of Computational Biology.* Department of Computer Science and Applied Mathematics, Weizmann Institute of Science. Web. 2010. <http://132.77.150.113/pubs/mir07/mir07_prediction.html>.

Persson, Helena, Anders Kvist, Johan Vallon-Christersson, Patrik Medstrand, Ake Borg, and Carlos Rovira. "The Non-coding RNA of the Multidrug Resistance-linked Vault Particle Encodes Multiple Regulatory Small RNAs." *Nature Cell Biology* 11.10 (2009): 1268-271. Print.

Peter, ME. "Targeting of MRNAs by Multiple MiRNAs: the next Step." *Oncogene* 29 (2010): 2161-164. Print.

"Really New Advances." *Economist* 383.8533 (2007): 87-89. Print.

Scheffer, George L., Peter L.J. Wijngaard, Marcel J. Flens, Miguel A. Izquierdo, Marilyn L. Slovak, Herbert M. Pinedo, Chris J.L.M. Meijer, Hans C. Clevers, and Rik J. Scheper. "The Drug Resistance-related Protein LRP Is the Human Major Vault Protein." *Nature Medicine* 1.6 (1995): 578-82. Print.

Van Zon, Arend, Marieke H. Mossink, Martijn Schoester, George L. Scheffer, Rik J. Scheper, Pieter Sonneveld, and Erik A.C. Wiemer. "Multiple Human Vault RNAs." *The Journal of Biological Chemistry* 276.October 5 (2001): 37715-7721. Print.

Zuker, Michael, and Patrick Stiegler. "Optimal Computer Folding of Large RNA Sequences Using Thermodynamics and Auxiliary Information." *Nucleic Acids Research* 9.1 (1981): 133-48. Print.